**Iranian Journal of Animal Science Research**
Homepage: http://ijasr.um.ac.ir

**Research Paper**
**Vol. 15, No. 1, Spring 2023, p. 107-121**

# Study of Differential Genes Expression of Uterine tissue Related to EggShell Using Transcriptomics data

**Milad Gholami Tahoone [1]*, Hossein Moradi Shahrebabak [2]**

Received: 28-11-2021
Revised: 13-06-2022
Accepted: 25-06-2022
Available Online: 25-06-2022

**How to cite this article:**
Gholami Tahoone, M., & Moradi Shahrebabak, H.(2023). Study of differential genes expression of uterine tissue related to eggshell using transcriptomics data. *Iranian Journal of Animal Science Research*, *15*(1), 107-121.
DOI: 10.22067/ijasr.2022.73912.1055

**Introduction:** Due to the high importance of egg shell quality in terms of hardness, extensive transportation with high volume of this product between different countries, the importance of shell quality in terms of protecting the internal content from extra-shell contamination and also one of the ways to improve traits, relying on The genetic structure of a trait is the application of genetic selection. Therefore, the aim this study is to use transcriptome data obtained from uterine tissue of two groups of laying hens with hard and weak eggshell, in each of the three stages of the calcification period (initiation, growth and termination). Gene expression profile was determined and differential expression of genes is analyzed so that different index genes can be expressed and their ontology can be studied to help the results. As a result, the genetic structure of the trait and the list of index genes have been regularly completed, and with the information obtained from the present study, the correct and practical genetic selection can be made in laying hen breeds.

**Materials and Methods:** The aim of this study was to investigate the gene expression profile between two groups of chickens with hard and weak eggshells in three stages of calcification period (initiation, growth and termination), to identify significant index genes with different expression and analysis. Ontology and the paths involved are related to them. Therefore, in this study, transcripts (total mRNA sequences) of 18 samples of uterine tissue including three replicates in each stage of each group were prepared. The process of converting the initial format of readings was done using SRAtoolkit software, measuring the quality of readings with fastQC software, editing readings with Trimmomatic software and re-measuring the quality of readings after editing with fastQC software; Also, the alignment and mapping of the readings were done using Hisat2 software and format conversion and sorting of results was performed using Samtools software. Finally, the differential expression analysis of genes was performed using Cufflinks software package (Cufflink, Cuffmerge and Cuffdiff software) and Then, considering p-value <0.00025 and the criterion for comparing the value of differential gene expression, log2 (fold_change) less than -4 and greater than +4 were identified as the most significant genes with different and significant expression. Histogram of the results was performed using CummeRbund package in R software environment and ontological analysis was performed using DAVID and Ensembl databases.

**Results and Discussion:** According to the results of gene expression profile, 28833 genes were identified on the transcript of the samples. Finally, by analyzing the differential expression of genes, 1338 genes at the initiation stage, 81 genes at the growth stage and 190 genes at the termination stage were identified with significantly different expressions, that Considering log2(fold_change) > +4 and log2(fold_change) < -4, 51 genes at the initiation stage, two genes at the growth stage and four genes at the termination stage had different and significant expression. The largest difference in gene expression between the two groups at the initiation stage,

1-Graduted M.Sc. Student, Department of Animal Science, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran.
2-Associate professor, Department of Animal Science, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran.
*Corresponding Author's Email: gholami.milad@ut.ac.ir

ENSGALG0000000044418 gene with log2 (fold_change) equal to 8/63 and at the growth stage, ENSGALG0000000049618 gene with log2 (fold_change) equal to 6/82 and at the termination stage, ENSGALG0000000049618 gene with log2 (fold_change) equal to 5/14 was observed. Ontological analysis of index genes showed that they are mainly involved in protein binding activities (ENSGAL000000444418, NMRAL1, SLITRK4, CHL1, ENSGAL00000009041, PHACTR3 and CORIN genes), DNA transcription regulation (LHX1, TFAP2A, RUNX3, ETV1, NR4A3 and RUNX1 genes), immune response (ENSGAL0000000043996, TMEM117, AvBD9, GCNT3, ENSGAL00000046947 and ENSGAL00000051617 genes), fat metabolism (PNPLA3, NR4A3, CD36, FFAR4 and DGKB genes) and calcium ion binding (ANXA10, CAPN8, CDH18 and DGKB genes), respectively.

**Conclusion:** In the differential expression analysis of genes, considering log2 (fold_change) more than +4 and less than -4, in the initial stage, 51 genes, in the growth stage, two genes and in the termination stage, four genes were identified. The highest differences in gene expression between the two groups were observed in the initial stage, ENSGALG00000044418 and PDZK1IP1 genes, in the growth stage, ENSGALG00000049618, ENSGALG00000048945 and TMEM63C genes and in the termination stage, ENSGALG000000004400 and ENSG26 genes; The results of ontological analysis of index genes showed that they are mainly involved in protein binding, DNA transcription regulation, immune response, lipid metabolism and calcium ion binding, respectively.

**Keywords:** Calcification, Differential gene expression, Eggshell, Gene ontology, Transcriptome

**نشریه پژوهشهای علوم دامی ایران**
Homepage: https://ijasr.um.ac.ir/

**مقاله پژوهشی**
**جلد۱۵، شماره۱، بهار۱٤۰۲، ص ۱۲۱-۱۰۷**

# مطالعه بیان افتراقی ژن‌های بافت رحم مرتبط با کیفیت پوسته تخم مرغ با استفاده از داده‌های ترانسکریپتومیکس

**میلاد غلامی طاحونه[۱]، حسین مرادی شهربابک[۲]**

تاریخ دریافت: ۱۴۰۰/۰۹/۰۷
تاریخ بازنگری: ۱۴۰۱/۰۳/۲۳
تاریخ پذیرش: ۱۴۰۱/۰۴/۰۴

## چکیده

با توجه به نقش کیفیت مطلوب پوسته تخم مرغ در مواردی همچون حمل‌ونقل و تبادلات گازی محتوای داخل تخم مرغ با بیرون از آن و نیز یکی از روش‌های بهبود صفات، با تکیه بر ساختار ژنتیکی صفت، اعمال انتخاب ژنتیکی است؛ در این پژوهش با هدف مطالعه پروفایل بیان ژن و تعیین ژن‌های شاخص با بیان متفاوت، داده‌های خام توالی ترانسکریپتوم دو گروه مرغ با پوسته تخم سخت و ضعیف در سه مرحله از دوره آهکی شدن پوسته به‌ترتیب شامل مراحل شروع، رشد و خاتمه مورد بررسی قرار گرفت. ترانسکریپتوم (توالی کل mRNA) ۱۸ نمونه از بافت رحم با سه تکرار در هر مرحله از هر گروه از پایگاه داده NCBI اخذ گردید و سپس به‌ترتیب مراحل پیش‌پردازش خوانش‌ها با نرم‌افزارهای SRAtookit، fastQC و Trimmomatic، هم‌ردیفی و مکان‌یابی خوانش‌ها با نرم‌افزارهای Hisat2 و Samtools و سپس آنالیز بیان و بیان افتراقی با پکیج نرم‌افزاری Cufflinks انجام شد، در انتها آنالیز هستی‌شناسی ژن‌های شاخص در دو پایگاه داده DAVID و Ensemble مورد بررسی قرار گرفت. در نتیجه آنالیز بیان افتراقی، ۲۸۸۳۳ ژن بر روی ترانسکریپتوم نمونه‌ها شناسایی شد که در نهایت، ۱۳۳۸ ژن در مرحله شروع، ۸۱ ژن در مرحله رشد و ۱۹۰ ژن در مرحله خاتمه با بیان متفاوت و معنی‌دار شناسایی شد (P<0.00025). با در نظر گرفتن 4 <log2(fold_change) و - <log2(fold_change) 4، ۵۱ ژن در مرحله شروع، دو ژن در مرحله رشد و چهار ژن در مرحله خاتمه به‌عنوان شاخص‌ترین ژن‌ها با بیان متفاوت و معنی‌دار مشخص شدند. آنالیز هستی‌شناسی ژن‌های شاخص نشان داد، به‌طور عمده به‌ترتیب در فعالیت‌های اتصال پروتئین، تنظیم رونویسیDNA، پاسخ ایمنی، متابولیسم چربی و اتصال یون کلسیم نقش دارند.

**واژه‌های کلیدی:** آهکی شدن پوسته، بیان افتراقی ژن، پوسته تخم مرغ، ترانسکریپتوم، هستی‌شناسی ژن

## مقدمه

طبق آمار سازمان خواربار و کشاورزی ملل متحد (FAO) در سال ۲۰۲۰ میلادی در مجموع در جهان ۸۶/۶ میلیون تن تخم‌مرغ تولید شده‌است که کشورهای چین، سرزمین مرکزی چین و آمریکا به‌ترتیب با تولید ۳۰/۲، ۲۹/۸ و ۶/۶ میلیون تن رتبه های اول تا سوم را داشته‌اند که ایران با رتبه ۲۰ جهان، تولیدی بیش از ۷۵۵ هزار تن را ثبت نمودهاست؛ طبق آمار همین سازمان، در سال ۲۰۲۰ کشورهای نیوزلند، ترکیه و لهستان به‌ترتیب با میزان ۴۱۵، ۲۱۷ و ۱۹۸ هزار تن، بیشترین صادرات تخم‌مرغ با پوسته را در کارنامه خود ثبت کردند

۱- دانش آموخته کارشناسی ارشد، گروه علوم دامی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران.
۲-دانشیار، گروه علوم دامی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران.
(*- نویسنده مسئول: Email: gholami.milad@ut.ac.ir)

DOI: 10.22067/ijasr.2022.73912.1055

و ایران در سال ۲۰۲۰ میلادی توانسته ۱/۷ هزار تن صادرات انجام‌دهد (FAO,2020). طبق آخرین آمار منتشر شده جهانی، می‌توان نتیجه گرفت صنعت تولید تخم‌مرغ خوراکی، صنعتی با میزان تولید محصول زیاد و نیز نیازمند حمل‌ونقل اعم از صادرات و واردات است؛ اما به‌دلیل حساسیت پوسته تخم‌مرغ نسبت به نیروی خارجی و نیز با توجه به نزدیک شدن تولید تخم‌مرغ به حد بیولوژیکی یک تخم در روز، انجام گرفتن بهینه دوره آهکی شدن پوسته و نیز ایجاد تخم‌مرغ با ضخامت پوسته مطلوب، امری مؤثر بر اقتصاد این صنعت و نیز بازارپسندی محصول می‌باشد؛ از سوی دیگر، نازکی بیش از حد پوسته، خود عاملی بر افزایش آسیب‌پذیری محتوای درون تخم مرغ نسبت به عوامل آلوده‌کننده خارجی اعم از نفوذ گازها و مواد شیمیایی می‌شود.

بیان ژن مکانیسمی حساب‌شده، پیچیده و تنظیم‌گر است که باعث می‌شود، ژنوتیپ به‌صورت فنوتیپ ظاهر شود (Banabazi et al., 2012). در واقع کدهای ژنتیکی که در مولکول DNA ذخیره شده‌اند، به‌وسیله بیان ژن‌ها تفسیر می شوند و خصوصیات و نحوه بیان ژن باعث به‌وجود آمدن فنوتیپ در ارگانیسم خواهدشد (Banabazi et al., 2012). مطالعات بیان ژن به تولید حجم زیادی از داده‌ها در این زمینه منجر شده‌است که در میان انواع متعدد داده‌های بیان ژن، داده‌های RNA-seq به سبب شامل شدن توالی نواحی رونویسی‌شده و پوشش دادن همه این نواحی دارای مزیت زیادی نسبت به سایر انواع داده‌ها هستند؛ در سال‌های اخیر امکانات رایانه‌ای آنالیز این حجم از داده‌ها را در زمان و با هزینه کم امکان‌پذیر ساخته است. از مزایای استفاده از داده‌های RNA-seq، شناسایی ژن‌های جدید است (Gan et al., 2010; Sultan et al., 2008). اساس روش تولید داده‌های RNA-seq، بُرش کل ترانسکریپتوم (mRNA) به قطعات کوچک و سپس توالی‌یابی همه این قطعات است (Haas and Zody, 2010; Marguerat and Bahler, 2010). از این رو، از روی تعداد خوانش‌های مکان‌یابی‌شده برای یک ناحیه خاص، می‌توان به یک معیار دقیق از میزان رونویسی آن ناحیه رسید (Ekblom and Galindo, 2010; Wilhem and Landry,2009). شاخص بیان ژن یا به عبارت دیگر، معیار بیان ژن در این روش، محاسبه تعداد خوانش یا قطعه در حد کیلوباز به‌ازای هر یک میلیون خوانش مکان‌یابی‌شده می‌باشد (Mortazavi et al., 2008)، که این شاخص برای طول ژن و تعداد کل خوانش‌های مکان‌یابی‌شده در هر نمونه تصحیح شده و مقادیر برآورد شده برای ژن‌های مختلف با هر طول و تعداد خوانش قابل مقایسه و آزمون می‌باشد.

با توجه به اهمیت بالای صفت کیفیت پوسته تخم‌مرغ از نظر سختی که در حد مطلوب باشد و نیز یکی از روش‌های بهبود صفات، با تکیه بر ساختار ژنتیکی صفت، اعمال انتخاب ژنتیکی است؛ پس در این پژوهش، هدف بر آن است که با استفاده از داده‌های ترانسکریپتوم حاصل از بافت رحم دو گروه از مرغ‌های تخم‌گذار با پوسته‌تخم سخت و ضعیف، در هر یک از سه مرحله دوره آهکی شدن پوسته (شروع، رشد و خاتمه)، پروفایل بیان ژن‌ها تعیین و آنالیز بیان افتراقی ژن‌ها انجام شود، تا در نهایت، بتوان ژن‌های شاخص متفاوت بیان‌شده و نیز هستی شناسی آن‌ها مورد مطالعه قرار گیرد تا به کمک نتایج حاصل، ساختار ژنتیکی مربوط به صفت کیفیت پوسته تخم‌مرغ، هر چه بیشتر معین گردد.

## مواد و روش‌ها

### اخذ داده از پایگاه اطلاعاتی

داده‌های ترانسکریپتوم حاصل از ۱۸ نمونه بافت رحم در دو گروه با پوسته تخم‌مرغ سخت و ضعیف در سه مرحله متوالی شروع، رشد و خاتمه دوره آهکی شدن پوسته در مرغ‌های نژادهای- لاین- قهوه‌ای در سن ۴۲ هفتگی است که توسط دانشکده تغذیه و علوم تغذیه حیوانات دانشگاه کشاورزی هواجونگ چین به‌صورت کتابخانه ترانسکریپتوم با روش سانگر با استفاده از تکنیک ایلومینا توالی‌یابی و تهیه شده‌است. ۱۸ نمونه بافت مربوط به دو گروه و هر گروه شامل سه زیر گروه (سه مرحله آهکی شدن پوسته) بوده که هر زیر گروه شامل سه قطعه مرغ در دوره تخم‌گذاری بوده است. خوانش‌های حاصل از توالی به‌صورت جفتی (رفت و برگشت) با طول ۱۵۰-۵۰ جفت باز بودند. داده‌های خام از تاریخ ششم مارس سال ۲۰۱۹ در پایگاه اطلاعاتی NCBI با فرمت sra به‌صورت رایگان از طریق لینک زیر در دسترس است:

https://www.ncbi.nlm.nih.gov/sra/SRR8692161

پس از اخذ ۱۸ نمونه داده خام توالی ترانسکریپتوم از پایگاه داده NCBI، تمامی مراحل آماده‌سازی و تجزیه تحلیل آن‌ها در محیط گرافیکی سیستم عامل لینوکس اوبونتو سری ۲۱/۰۴ نسخه ۶۴ بیت بر روی کامپیوتر شخصی (تعداد هسته: هفت عدد، میزان رم: هشت گیگ و میزان حافظه: یک ترابایت) غیر متصل به سرور به‌ترتیب مراحل زیر انجام شد.

### تبدیل فرمت خوانش‌ها

به‌علت این که پکیج‌های نرم‌افزاری مورد استفاده در ادامه مسیر آنالیز، توانایی خوانش داده‌های ترانسکریپتومیکس با فرمت fastq را دارند، تبدیل فرمت sra به fastq با استفاده از نرم‌افزار sratoolkit نسخه ۲,۱۰,۲ و کد دستوری زیر انجام شد:

```
fastq-dump  --gzip  -O [output]  [sra_file]
```

### سنجش کیفیت خوانش‌ها

سنجش کیفیت با استفاده از نرم‌افزار fastQC نسخه ۰/۱۱/۹ در

محیط جاوا با کمک منوی اصـــلی، برای هر خوانش به‌طور جداگانه با ۱۰ آزمون مختلف انجام شـد (Andrews, 2010). نرم‌افزار نتیجه هر آزمون را جداگانه بررسـی نموده و وضـعیت قبول، هشـدار یا رد برای آن‌ها را به‌ترتیب با رنگ‌های سـبز، نارنجی و قرمز در کنار عنوان هر آزمون در پنل سمت چپ نرم‌افزار نشان می‌دهد. در نهایت، براساس نتیجه کیفیت همه نوکلئوتیدها، راجع به کیفیت طول کل خوانش تصمیم‌گیری می‌شود.

## ویرایش خوانش‌ها

در صـورتی که داده‌های جفتی (رفت و برگشـت) نیاز به ویرایش داشـته باشـند، نرم‌افزار Trimmomatic نسـخه ۰/۳۹ برای داده‌های توالی‌یابی نسل جدید شرکت ایلومینا بهینه‌سازی شده‌است (Bolger et al., 2014). وظایف این نرم‌افزار شـامل حذف آداپتورها و حذف یا ویرایش خوانش‌های بی‌کیفیت اسـت، که فهرسـت آداپتورهای به‌کار رفته در زمان توالی‌یابی با توجه به دسـتگاه مورد اسـتفاده به‌صـورت پیش‌فرض در برنامه قرار دارد. در این پژوهش، با در نظر گرفتن حداقل طول خوانش‌ها ۵۰ جفت باز، برای ویرایش خوانش‌ها از کد دستوری زیر استفاده شد:

```
Trimmomatic PE -threads 7 [input_data1.fq] [input_data2.fq] [clean_read1.fq] [trash_read1.fq] [clean_read2.fq] [trash_read2.fq] TRAILING:20 SLIDINGWINDOW:6:20 MINLEN:25
```

## سنجش مجدد کیفیت خوانش‌ها

پس از ویرایش، به جهت اطمینان از صـحت نتیجه آن، بار دیگر کیفیت خوانش‌ها با اسـتفاده از نرم‌افزار fastQC مورد سـنجش قرار می‌گیرد تا بتوان با توجه به نتیجه نهایی ویرایش، برای ادامه دادن مسیر تصمیم گیری نمود.

## تشکیل ترانسکریپتوم

از اهداف مهم آنالیز داده‌های ترانسکریپتوم، بازسازی مجموعه‌ای کامل از ترانسـکریپت ژن‌ها و نیز تخمین میزان سـطوح بیان تمام ترانسـکریپت‌هاسـت. برای محاسـبه سـطح بیان ترانسـکریپت‌ها، مکان‌یابی خوانش‌ها بر روی ژنوم مرجع اسـت که مکان دقیق خوانش‌ها را مشـخص می‌کند. برای انجام این مرحله، علاوه‌بر خوانش‌های ویرایش شـده، نیاز به داده‌های ژنوم مرجع و اطلاعات حاشیه‌نویسی است که آخرین نسخه به‌روز شده ژنوم مرجع مرغ به همراه فایل حاشیه‌نویسی مربوطه، از بانک اطلاعاتی Ensemble به آدرس‌های زیر دریافت شد:

http://ftp.ensembl.org/pub/release-104/fasta/gallus_gallus/dna/
http://ftp.ensembl.org/pub/release-104/gtf/gallus_gallus/

## هم‌ردیفی و مکان‌یابی خوانش‌ها بر روی ژنوم مرجع

در اسـتراتژی مکان‌یابی و هم‌ردیفی خوانش‌ها، ابتدا لازم اسـت ژنوم مرجع نمایه یا ایندکس شود تا نمایه کاملی از اطلاعات موجود در ژنوم مرجع حاصل شود تا در زمان مکان‌یابی بتوان بیشترین هم‌ردیفی را حاصـل نمود. بسـیاری از نرم‌افزارهای هم‌ردیف‌کننده، پیش از نقشـه یابی، ژنوم مرجع را ایندکس کرده و سـپس از ژنوم تازه ایندکس شده استفاده می‌کنند. برای ایندکس‌کردن ژنوم مرجع توسط نرم‌افزار Hisat2 نسخه ۲،۲،۱ از کد دستوری زیر استفاده شد:

```
Hisat2-build [genome-reference.fa] [output-directory/index-genome]
```

پس از ایندکس کردن ژنوم مرجع، هشـت فایل تولید شـد که به‌منظور افزایش میزان دقت و صحت نقشه‌یابی، علاوه‌بر این هشت فایل، از ژنوم مرجع و فایل حاشـیه‌نویسـی آن نیز برای هم‌ردیفی و نقشه‌یابی استفاده می شود. هم‌ردیفی خوانش‌های کوتاه بر روی ژنوم مرجع نیز با اسـتفاده از نرم‌افزار Hisat2 انجام‌شـد (Kim et al., 2015). ویژگی این نرم‌افزار در این است که با تأمین اطلاعات مربوط به ژن‌ها و انواع ترانسـکریپت آن‌ها، پیش آگاهی از محل اتصـال برش‌ها و در نظر گرفتن الگوهای جایگزین برای اتصـال برش‌ها، خوانش‌های مربوط به اتصال برش‌ها که بخشی از توالی آن‌ها با توالی اگزون قبل و بخشی دیگر با توالی اگزون‌های بعدی هم‌ردیف می‌گردد را با سـرعت و دقت بالاتری از نواسـمبل و مکان‌یابی می‌کند (Bahrami, 2020). برای هم‌ردیفی و مکان‌یابی خوانش‌ها بر روی ژنوم مرجع مرغ از کد دستوری زیر استفاده شد:

```
hisat2 -p 7 --dta-cufflinks --very-sensitive -x [index-genome] input-1.fq input-2.fq -S output.sam
```

برای انجام ادامه آنالیز، لازم اسـت که خروجی SAM مربوط به خوانش‌ها، به‌فرمت BAM که حاوی نتایج مکان‌یابی و پوشـش هر خوانش بر روی ژنوم می‌باشد، تبدیل شود. این تبدیل فرمت با استفاده از پکیج نرم‌افزاری samtools با کد دستوری زیر انجام شد:

```
samtools view -b -S -o [output-file.bam] [input-file.sam]
```

در ادامه لازم اسـت فایل‌های BAM تولیدشـده بر اسـاس کروموزوم‌ها به‌ترتیب مرتب شوند و سپس ایندکس از آن‌ها تهیه شود تا بتوان نمایه‌ای از اطلاعات به‌ترتیب شـماره کروموزوم حاصـل نمود، که با کد دستوری‌های زیر انجام شدند:

```
samtools sort [bam-file.bam] -o [output-sort.bam]
samtools index [sort.bam]
```

## سرهم کردن ترانسکریپتوم

به‌منظور هم‌ردیفی خوانش‌ها و سـرهم کردن ترنسـکریپتوم‌های موجود، از بسـته نرم‌افزار Cufflinks نسـخه ۲/۲/۱ اسـتفاده شـد (Github, 2017). در این بسته نرم‌افزاری دلیل اصلی سرهم کردن مجدد، در نظر گرفتن تمام ترانسـکریپت‌ها، ژن‌ها، اگزون‌ها و سـایر

ویژگی‌های مختص تیمار مورد بر رسی ا ست که این ویژگی‌های ژنی در ترانسکریپتوم مرجع موجود نبودند. کد دستوری زیر جهت هم‌ردیفی و سرهم کردن خوانش‌ها استفاده شد:

cufflinks -p 7 -g [reference-genome.gtf] -o [output-name] -L [output-name] [sort.bam]

## تلفیق ترانسکریپتوم‌ها

بعد از اتمام هم‌ردیفی و اسمبلی هرکدام از نمونه‌ها، ترانسکریپتوم‌های تشکیل شده از هر نمونه توسط نرم‌افزار cuffmerge با یکدیگر ادغام می‌شوند و یک فایل ترانسکریپتوم جدید مرجع ادغام‌شده برای تمامی تکرارهای هر نمونه حاصل می‌شود. لازم به ذکر ا ست، قبل از تلفیق ترانسکریپتوم، ترانسکریپتوم‌های هر گروه در قالب یک فایل با فرمت متن ذخیره شده‌اند. برای ادغام از کد دستوری زیر استفاده شد:

cuffmerge -p 7 -o [ortput-merge] -g [reference-genome.gtf] [assembly-grf-list.txt]

## محاسبه بیان افتراقی ژن

برای تست آماری میزان بیان ژن بین دو نمو نه، از نرم‌افزار cuffdiff استفاده شد. این نرم‌افزار ژن‌هایی که از نظر آماری هر گونه تفاوت معنی‌داری در بیان را نشان می‌دهند، شناسایی می‌کند. با استفاده از کد دستوری زیر، ژن‌های با بیان متفاوت شناسایی شدند:

cuffdiff -p 7 --library-norm-method classic-fpkm -o [output-file] -L [input1,input2] [merge.gtf] [input1-sort.bam] [input2-sort.bam]

## مصورسازی نتایج

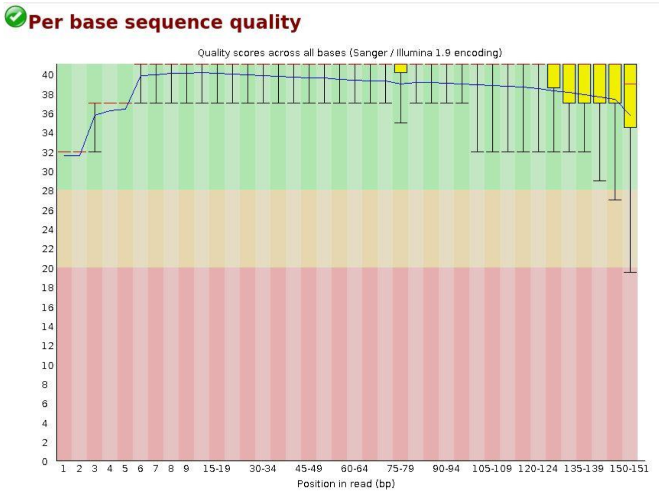
فایل خروجی آنالیز بیان افتراقی ژن‌ها با استفاده از Excel قابل مشاهده است، ولی برای مصورسازی نتایج در قالب پلات و گراف‌های قابل فهم‌تر، از پکیج CummeRbund در محیط نرم‌افزار R استفاده شد.

## آنالیز هستی‌شناسی ژن

در هستی‌شناسی ژن، داده‌های مورد نظر طبقه‌بندی و ارتباط آن‌ها با هم مطالعه می شوند. در این پژوهش، آنالیز هستی شناسی ژن‌های معنی‌دار با بیان متفاوت، به‌وسیله پایگاه‌های مربوط به هستی شناسی DAVID (https://david.ncifcrf.gov/) و Ensemble (https://www.ensembl.org/Chicken/) انجام شد.

## نتایج و بحث

پس از تبدیل فرمت داده‌ها و انجام سنجش کنترل کیفیت، طبق گزارش حاوی آماره‌های پایه، توالی‌های بین ۱۹ تا ۲۳ میلیون خوانش کوتاه جفتی (از هر دو انتهای توالی‌یابی شده) با طول ۱۵۰-۵۰ جفت باز برای نمونه ها به‌دست آمد. در آزمون کنترل کیفیت تک نوکلئوتیدی، محور افقی نمودار بیانگر موقعیت هر جفت باز بر روی خوانش و محور عمودی بیانگر نمره فِرِد (phred score) است که نمونه با نمره بالاتر از ۲۸ (منطقه سبز) به‌عنوان نمونه با کیفیت بالا تلقی می‌شود که طبق گراف مربوط به نمونه مورد نظر، کیفیت نمونه خوب می‌باشد (شکل ۱).

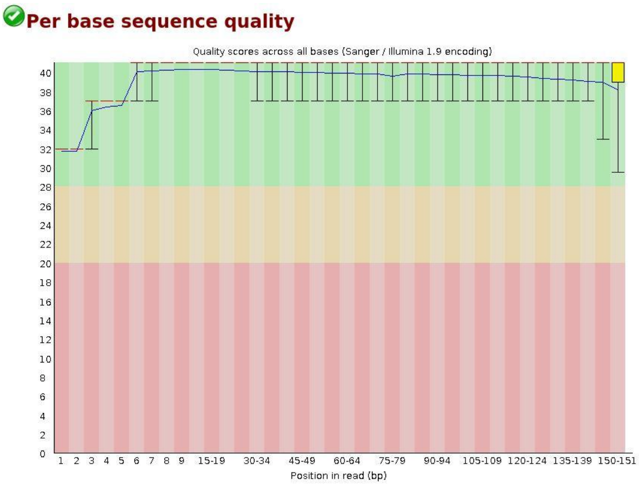



**شکل ۱**- آزمون کنترل کیفیت تک نوکلئوتیدی مربوط به نمونه "SRR8692161"

**Figure 1**- per base sequence quality related to "SRR8692161" sample.

با در نظر گرفتن حداقل طول خوانش ۵۰ جفت باز و سنجش کنترل کیفیت پس از ویرایش مشخص گردید که ۱۸ تا ۲۲ میلیون خوانش با کیفیت مناسب برای نمونه باقی ماند و کمتر از ۱۰ درصد از کل خوانش‌ها کنار گذاشته شد و طبق آزمون کنترل کیفیت تک نوکلئوتیدی، نمره فِرِد نوکلئوتید های انتهای توالی خوانش افزایش یافته و کیفیت نمونه ها نسبت به قبل افزایش یافت و توالی

خوانش‌های با کیفیت بالا (منطقه سبز رنگ) نگهداری شدند (شکل ۲).




**شکل ۲**- آزمون کنترل کیفیت تک نوکلئوتیدی پس از ویرایش مربوط به نمونه "SRR8692161"

**Figure 2**- per base sequence quality after trimming related to "SRR8692161" sample

طبق اطلاعات خروجی در انتهای آنالیز هم‌ردیفی و مکان یابی خوانش‌ها بر روی ژنوم مرجع، میانگین درصد مکان‌یابی خوانش‌ها بیش از ۸۰ درصد و خوانش‌های با محل هم‌ردیفی چندگانه کمتر از یک درصد بودند و با تشکیل پروفایل بیان ژن برای هر یک از زیر گروه‌های حاوی نمونه در هر یک از مراحل دوره آهکی شدن پوسته و نیز آنالیز بیان افتراقی بین زیر گروه‌های مشابه در بین دو گروه نمونه حاصل از مرغ‌های با پوسته سخت و ضعیف ، در نهایت با تجزیه‌تحلیل نتایج خروجی، ۳۵۳۴۲ ژن و ۱۰۷۷۲۲ ایزوفرم شناسایی شد که از این میان ۱۳۳۸ ژن با بیان متفاوت در مرحله شروع، ۸۱ ژن در مرحله رشد و ۱۹۰ ژن در مرحله خاتمه مشخص شد ($p<0.00025$). در ادامه به جهت رسیدن به تصویر مناسبی از تغییرات بیان در مقایسه‌های دو به دو و شناسایی ژن‌های شاخص با بیان متفاوت، میزان معیار log2(fold_change)به‌گونه‌ای در نظر گرفته شده است که در هر سه مرحله، شاخص ترین ژن‌ها مورد شناسایی و بررسی قرار گیرند؛ از این رو، ژن‌های با log2(fold_change) بیشتر از ۴+ و کمتر از ۴- برای مطالعه انتخاب شدند که شامل ۵۱ ژن در مرحله شروع (جدول ۱)، دو ژن در مرحله رشد (جدول ۲) و چهار ژن در مرحله خاتمه (جدول ۳) می‌باشند.

نتایج جداول فوق نشان می‌دهند که در مرحله شروع، تعداد ژن‌های شاخص بیشتری با میزان معیار log2(fold_change)>+4 & log2(fold_change)<-4 نسبت به دو مرحله دیگر شناسایی شده‌اند، که خود می‌تواند بیانگر وجود تفاوت بیان بیشتر در سطح ژنوم بین دو زیر گروه در این مرحله باشد؛ همچنین در سه مرحله دوره آهکی شدن پوسته، تنها شناسه ژن XLOC_008910 ما بین دو مرحله رشد و خاتمه به‌عنوان ژن‌های شاخص با بیان متفاوت، مشترک شناسایی شده‌اند.

نمودار هیستوگرام ژن‌های شاخص مورد مطالعه بین دو گروه در هر یک از مراحل دوره آهکی شدن پوسته ترسیم گردید (شکل ۳، ۴ و ۵). در نمودار هیستوگرام، محور افقی شامل ژن‌های شاخص مورد مطالعه (log2(fold_change)>+4 & log2(fold_change)<-4) و محور عمودی بیانگر شاخص قطعه در حد کیلوباز به‌ازای هر یک میلیون خوانش مکان‌یابی‌شده به‌علاوه یک واحد (FPKM+1) است.

با توجه به ارتفاع ستون‌های مندرج در شکل ۳، به‌ترتیب ژن‌های، ENSGALG00000046947 و PNPLA3 در گروه با پوسته تخم سخت و تنها ژن NR4A3 در گروه با پوسته تخم ضعیف بیشترین بیان را داشته‌اند و نیز بیشترین تفاوت بیان مربوط به ژن ENSGALG00000044418 با میزان log2(fold_change) برابر با ۸/۶ است؛ ارتفاع ستون های مندرج در شکل ۴، ژن ENSGALG00000049618 در گروه با پوسته تخم سخت و ژن ENSGALG00000048945 در گروه با پوسته تخم ضعیف بیشترین بیان را داشته‌اند و نیز بیشترین تفاوت بیان مربوط به ژن ENSGALG00000049618 با میزان log2(fold_change) برابر با ۶/۸ است و ارتفاع ستون‌های مندرج در شکل ۵، بیان تمامی ژن‌های شاخص در مرحله خاتمه دوره آهکی شدن پوسته، در گروه با پوسته تخم ضعیف بیشتر از گروه دیگر بوده است و ژن MIR12247-1 بیشترین بیان را داشته است؛ همچنین در مقایسه بین دو گروه در مرحله خاتمه، ژن ENSGALG00000049618 با میزان log2(fold_change) برابر با ۵/۱ بیشترین تفاوت بیان را داشته‌اند.

آنالیز هستی‌شناسی ژن‌های شاخص مورد مطالعه در مرحله شروع (جدول ۴)، مرحله رشد (جدول ۵) و مرحله خاتمه (جدول ۶) بر اساس سه دسته بندی عمومی فرآیندهای زیستی، عملکرد مولکولی و اجزای

سلولی انجام و ارائه شده است.

**جدول ۱**- شناسه، نام یا کد، جایگاه و ارزش p ژن‌های شاخص متفاوت بیان شده در مرحله شروع دوره آهکی شدن پوسته

**Table 1**- Identifier, name or code, position and p-value of indicator genes differentially expressed in the initiation stage of the calcification period

| شناسه ژن<br>Gene_id | نام یا کد ژن<br>Name or gene code | جایگاه<br>Locus | ارزش p<br>p_value |
|---|---|---|---|
| XLOC_022317 | ENSGALG00000044418 | 5:14892237-14898052 | 0.00005 |
| XLOC_021085 | ENSGALG00000043996<br>ENSGALG00000044755<br>ENSGALG00000045876<br>ENSGALG00000047967<br>ENSGALG00000048572<br>ENSGALG00000052476 | 4:85657237-85900720 | 0.00005 |
| XLOC_016325 | ENSGALG00000052817 | 3:69232233-69353759 | 0.00005 |
| XLOC_009120 | LHX1 | 19:8535713-8540691 | 0.00005 |
| XLOC_019526 | ANXA10 | 4:24919747-24946202 | 0.00005 |
| XLOC_002605 | PNPLA3 | 1:69011399-69090996 | 0.00005 |
| XLOC_013171 | ENSGALG00000049241 | 22:76807-89630 | 0.00005 |
| XLOC_002772 | UPK1B | 1:81562062-81573670 | 0.00005 |
| XLOC_000247 | TMEM117 | 1:30224568-30431209 | 0.00005 |
| XLOC_010329 | ENSGALG00000015234<br>ENSGALG00000053346 | 2:108071663-108150364 | 0.00005 |
| XLOC_010003 | TFAP2A | 2:63284964-63301149 | 0.00005 |
| XLOC_002039 | CD36 | 1:11331665-11404784 | 0.00005 |
| XLOC_004693 | GCNT3 | 10:5852886-5887192 | 0.00005 |
| XLOC_016354 | - | 3:73133293-73138023 | 0.00005 |
| XLOC_012502 | MMP9 | 20:10572273-10576665 | 0.00005 |
| XLOC_016106 | - | 3:46416828-46464624 | 0.00005 |
| XLOC_019359 | SLITRK4 | 4:9895861-9908452 | 0.00005 |
| XLOC_003777 | - | 1:32040296-32044590 | 0.00005 |
| XLOC_005715 | CHL1 | 12:17603135-17708756 | 0.00005 |
| XLOC_017188 | ZC3H12D | 3:48201990-48215042 | 0.00005 |
| XLOC_011706 | - | 2:120693001-120714502 | 0.00005 |
| XLOC_010601 | ADCY8 | 2:140875169-140995082 | 0.00005 |
| XLOC_000699 | PNPLA3 | 1:69011399-69090996 | 0.00005 |
| XLOC_015702 | ENSGALG00000009041 | 3:8772208-8777996 | 0.00005 |
| XLOC_002973 | RUNX1 | 1:107353573-107513309 | 0.00005 |
| XLOC_012404 | PHACTR3 | 20:6735474-6830610 | 0.00005 |
| XLOC_021710 | ENSGALG00000008599 | 5:24851725-24885341 | 0.00005 |
| XLOC_018319 | ENSGALG00000046947 | 31:2413209-2413512 | 0.00005 |
| XLOC_010948 | DGKB | 2:27385425-27720368 | 0.00005 |
| XLOC_012126 | - | 2:79230058-79235013 | 0.00005 |
| XLOC_018508 | ENSGALG00000051617 | 31:2422484-2443234 | 0.00005 |
| XLOC_011589 | OPN5L1 | 2:106222092-106258854 | 0.00005 |
| XLOC_011878 | ADCY8 | 2:140875169-140995082 | 0.00005 |
| XLOC_013499 | RUNX3 | 23:2591780-2635428 | 0.00005 |
| XLOC_012132 | - | 2:79639930-79644453 | 0.00005 |
| XLOC_007199 | NMRAL1 | 14:13089511-13149359 | 0.00005 |
| XLOC_015815 | CAPN8 | 3:17803815-17833098 | 0.00005 |
| XLOC_025365 | PAPPA2 | 8:7156618-7232054 | 0.00005 |
| XLOC_010947 | ETV1 | 2:27271071-27337505 | 0.00005 |
| XLOC_013760 | ENSGALG00000051964 | 23:5829851-5834748 | 0.00005 |
| XLOC_003775 | - | 1:32027118-32032886 | 0.00005 |
| XLOC_011450 | NR4A3 | 2:89336025-89389150 | 0.00005 |
| XLOC_025534 | PDZK1IP1 | 8:22247457-22251895 | 0.0001 |
| XLOC_010076 | CDH18 | 2:74189948-74446617 | 0.0001 |
| XLOC_016657 | AvBD9 | 3:107430691-107459717 | 0.0001 |
| XLOC_019907 | CORIN | 4:66638160-66761091 | 0.00015 |
| XLOC_007726 | KREMEN1 | 15:8089944-8111256 | 0.0002 |
| XLOC_007047 | CYP3A5 | 14:4322582-4332428 | 0.0002 |
| XLOC_021886 | - | 5:41926379-42116783 | 0.00025 |

**جدول ۲**- شناسه، نام یا کد، جایگاه و ارزش p ژن‌های شاخص متفاوت بیان شده در مرحله رشد دوره آهکی شدن پوسته

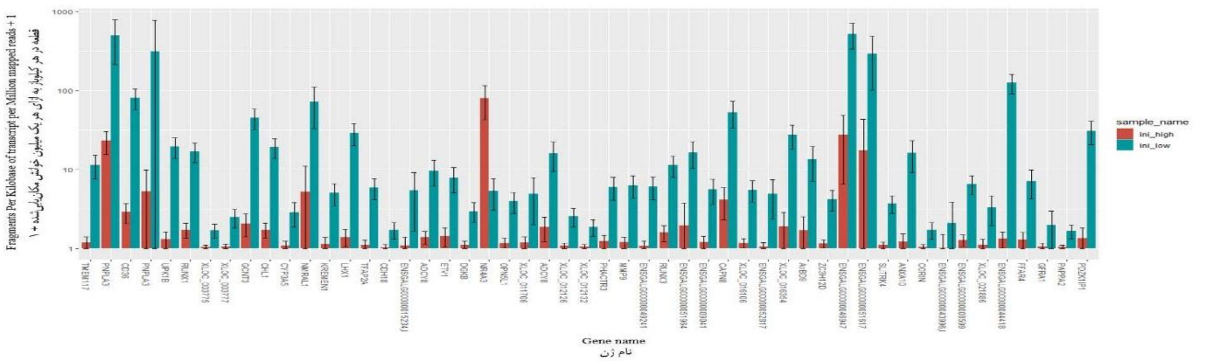
**Table 2**- Identifier, name or code, position and p-value of indicator genes differentially expressed in the growth stage of the calcification period

| شناسه ژن<br>Gene_id | نام یا کد ژن<br>Name or gene code | جایگاه<br>Locus | ارزش p<br>P_value |
|---|---|---|---|
| XLOC_008910 | ENSGALG00000049618 | 18:11326495-11332858 | 0.00005 |
| XLOC_021869 | ENSGALG00000048945 | 5:39323944-39354365 | 0.00005 |

**جدول ۳**- شناسه، نام یا کد، جایگاه و ارزش p ژن‌های شاخص متفاوت بیان شده در مرحله خاتمه دوره آهکی شدن پوسته

**Table 3-** Identifier, name or code, position and p-value of indicator genes differentially expressed in the termination stage of the calcification period

| شناسه ژن<br>Gene_id | نام یا کد ژن<br>Name or gene code | جایگاه<br>Locus | ارزش p<br>p_value |
|---|---|---|---|
| XLOC_008910 | ENSGALG00000049618 | 18:11326495-11332858 | 0.00005 |
| XLOC_007996 | ENSGALG00000054226 | 16:1702927-1709590 | 0.00005 |
| XLOC_014979 | HSPB9 | 27:7654019-7655569 | 0.00005 |
| XLOC_013278 | MIR12247-1 | 22:5177338-5230538 | 0.00005 |

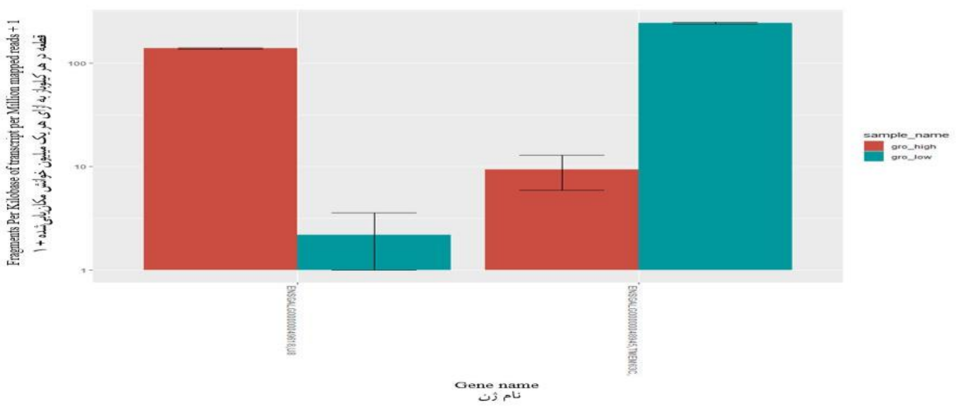



**شکل ۳**- نمودار هیستوگرام تفاوت بیان ژن‌های شاخص در مرحله شروع دوره آهکی شدن پوسته بین دو گروه

(ini_high: مرحله شروع با پوسته سخت و ini_low: مرحله شروع با پوسته ضعیف)

**Figure 3-** Histogram plot of differential indicator genes expression in the initiation stage of the calcification period between two groups

(ini_high: initiation stage with high eggshell, ini_low: initiation stage with low eggshell)

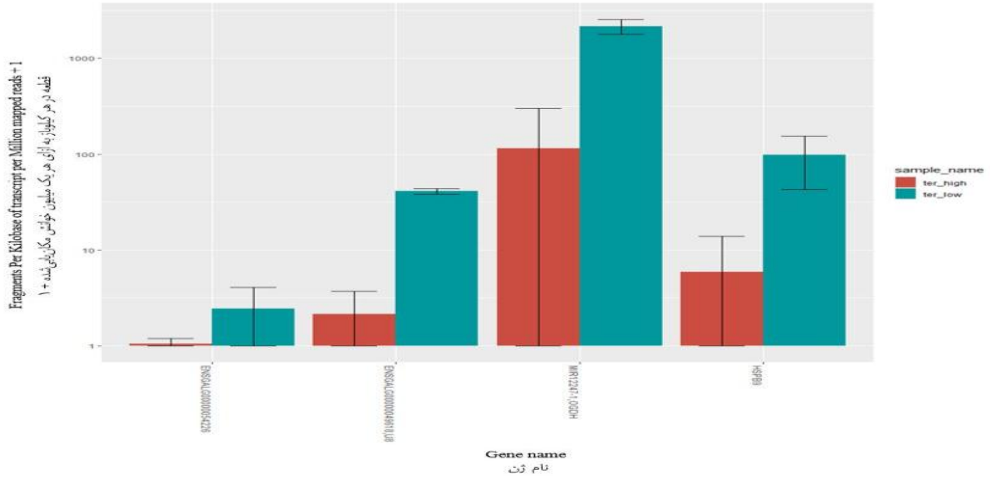



**شکل ٤**- نمودار هیستوگرام تفاوت بیان ژن‌های شاخص در مرحله رشد دوره آهکی شدن پوسته بین دو گروه

(gro_high: مرحله رشد با پوسته سخت و gro_low: مرحله رشد با پوسته ضعیف)

**Figure 4-** Histogram plot (Right) & Heatmap plot (Left) of differential indicator genes expression in the growth stage of the calcification period between two groups

(gro_high: growth stage with high eggshell, gro_low: growth stage with low eggshell)




**شکل ۵**- نمودار هیستوگرام تفاوت بیان ژن‌های شاخص در مرحله خاتمه دوره آهکی شدن پوسته بین دو گروه
(ter_high: مرحله خاتمه گروه با پوسته سخت و ter_low: مرحله خاتمه گروه با پوسته ضعیف)

**Figure 5**- Histogram plot of differential indicator genes expression in the termination stage of the calcification period between two groups
(ter_high: termination stage with high eggshell, ter_low: termination stage with low eggshell)

**جدول ٤**- آنالیز هستی‌شناسی ژن‌های شاخص متفاوت بیان شده در مرحله شروع دوره آهکی شدن پوسته

**Table 4**- Gene ontology analysis of indicator genes differentially expressed in the initiation stage of the calcification period

| نام ژن<br>Gene name | فرآیند زیستی<br>Biological process | اجزای سلولی<br>Cellular components | عملکرد مولکولی<br>Molecular function |
|---|---|---|---|
| ENSGALG00000044418 | تنظیم رگ‌زایی<br>Regulation of angiogenesis | فضای خارج سلولی<br>Extracellular space | اتصال به پروتئین<br>Protein binding |
| ENSGALG00000043996<br>ENSGALG00000044755<br>ENSGALG00000045876<br>ENSGALG00000047967<br>ENSGALG00000048572<br>ENSGALG00000052476 | ایمنی به‌واسطه سلول T، مسیر سیگنالینگ گیرنده سطح سلول<br>T cell-mediated immunity, A cell surface receptor signaling pathway | غشاء<br>membrane | – |
| LHX1 | توسعه رحم و تخمدان، توسعه و تمایز مخچه، تنظیم بیان ژن، تمایز سلولی، تمایز نورونی<br>Uterus and ovary development, Cerebellum development and differentiation, Regulation of gene expression, Cell differentiation, Neuronal differentiation | هسته، مجموعه پروتئین<br>Nucleus, Protein complex | فعالیت فاکتور رونویسی اتصالDNA، اتصالDNA<br>DNA binding transcription factor activity, DNA binding |
| ANXA10 | – | میتوکندری<br>Mitochondria | اتصال یون کلسیم، اتصال فسفولیپید وابسته به کلسیم<br>Calcium ion binding, Calcium-dependent phospholipid binding |
| PNPLA3 | فرآیند کاتابولیک و بیوسنتز تری گلیسیرید، فرآیند متابولیک لیپید، هموستاز چربی<br>Catabolic process and triglyceride biosynthesis, Lipid metabolic process, Lipid homeostasis | غشاء، سیتوپلاسم، قطره چربی<br>Membrane, Cytoplasm, Lipid droplet | فعالیت تری گلیسیرید لیپاز، فسفولیپازA2 و هیدرولاز<br>Triglyceride lipase, Phospholipase A2 and Hydrolase activity |
| UPK1B | تمایز سلول‌های اپیتلیال<br>Differentiation of epithelial cells | غشاء، غشای پلاسمایی آپیکال<br>Membrane, Apical plasma membrane | – |

| | | | |
|---|---|---|---|
| TMEM117 | مسیر سیگنال‌دهی آپوپتور در پاسخ به استرس شبکه آندوپلاسمی<br>Apoptotic signaling pathway in response to endoplasmic reticulum stress | غشاء، شبکه آندوپلاسمایی<br>Membrane, Endoplasmic reticulum | – |
| ENSGALG00000015234<br>ENSGALG00000053346 | انتقال غشاء کلرید<br>Chloride membrane transport | غشاء<br>Membrane | فعالیت انتقال‌دهنده غشاء کلرید<br>Membrane chloride transporter activity |
| TFAP2A | تنظیم رونویسی DNA، تنظیم تکثیر سلولی، توسعه ساختار آناتومی<br>Regulation of DNA transcription, Regulation of cell proliferation, Development of anatomical structure | هسته، دستگاه گلژی، سانتروزوم<br>Nucleus, Golgi apparatus, Centrosome | فعالیت فاکتور رونویسی متصل بهDNA، فعال‌کننده و سرکوب‌کننده رونویسی<br>DNA binding transcription factor activity, Transcription activator and repressor |
| RUNX3 | فسفوریلاسیون پروتئین، تنظیم تمایز سلولی<br>Protein phosphorylation, Regulation of cell differentiation | هسته، سیتوپلاسم<br>Nucleus, cytoplasm | اتصال فاکتور رونویسی ATP، اتصال DNA<br>ATP binding of transcription factor, DNA binding |
| NMRAL1 | – | هسته، نوکلئوپلاسم<br>Nucleus, Nucleoplasm | اتصال پروتئینی یکسان<br>Identical protein binding |
| CAPN8 | هضم، پروتئولیز<br>Digestion, Proteolysis | – | اتصال یون کلسیم، فعالیت اندوپپتیداز<br>Calcium ion binding, Endopeptidase activity |
| GFRA1 | توسعه سیستم عصبی<br>Development of the nervous system | غشاء، اگزوزوم خارج سلولی<br>Membrane, Extracellular exosome | فعالیت گیرنده<br>Receiver activity |
| PAPPA2 | مورفوژنز استخوان، پروتئولیز<br>Bone morphogenesis, Proteolysis | – | اتصال یون روی<br>Zinc ion binding |
| ETV1 | هدایت آکسون، تمایز سلولی، توسعه اندام عضلانی<br>Axon guidance, Cell differentiation, Muscle organ development | هسته<br>Nucleus | اتصال فعال‌کننده رونویسی متصل به DNA<br>DNA bound transcriptional activator |
| NR4A3 | تنظیم رونویسی الگو DNA، تنفس سلولی، تمایز سلول‌های چربی، هموستاز انرژی، تنظیم چرخه سلولی<br>DNA template transcription regulation, Cellular respiration, Adipocyte differentiation, Energy homeostasis, Cell cycle regulation | هسته، مجموعه فاکتور رونویسی<br>Nucleus, Transcription factor complex | اتصال فعال‌کننده رونویسی اتصال به DNA، اتصال یون روی، اتصال گیرنده هورمون استروئیدی<br>Transcription activator binding, DNA binding, Zinc ion binding, Steroid hormone receptor binding |
| PDZK1IP1 | – | غشاء، اگزوزوم خارج سلولی<br>Membrane, Extracellular exosome | – |
| CDH18 | چسبندگی سلول‌های هموفیلی از طریق مولکول‌های چسبندگی غشای پلاسمایی<br>Adhesion of hemophilia cells through plasma membrane adhesion molecules | غشاء<br>Membrane | اتصال یون کلسیم<br>Calcium ion binding |
| AvBD9 | پاسخ دفاعی به باکتری گرم منفی و قارچ، شیمیوتاکسی سلولی<br>Defense response to Gram-negative bacteria and fungi, Cellular chemotaxis | سیتوپلاسم، فضای خارج سلولی<br>Cytoplasm, Extracellular space | اتصال شیمیایی جذب‌کننده، اتصال گیرنده‌های شیمیوکاین<br>Chemoattractant binding, Binding of chemokine receptors |
| CD36 | انتقال تری گلیسیرید، فاگوسیتوز، تنظیم ذخیره کلسترول و چربی، تنظیم هموستاز انرژی، پاسخ به اسید لینولئیک<br>Triglyceride transfer, Phagocytosis, | دستگاه گلژی، غشاء، فضای خارج سلولی<br>Golgi apparatus, | فعالیت گیرنده لیپوپروتئین با چگالی کم، اتصال چربی<br>Low-density lipoprotein |

| | | | |
|---|---|---|---|
| | Regulation of cholesterol and fat storage, Regulation of energy homeostasis, .Response to linoleic acid | Membrane, Extracellular space | Receptor activity, Lipid binding |
| GCNT3 | مورفوژنز بافت و کلیه، تولید ایمونوگلوبولین در بافت مخاط<br>Tissue and kidney morphogenesis, Immunoglobulin production in mucous tissue | غشاء، غشای گلژی<br>Membrane, Golgi membrane | فعالیت ترانسفراز انتقال گروه‌های گلیکوزیل<br>Transferase activity of glycosyl group transfer |
| MMP9 | توسعه سیستم اسکلتی، تنظیم اتصال DNA، تنظیم مثبت فسفوریلاسیون پروتئین<br>Development of the skeletal system, Regulation of DNA binding, Positive regulation of protein phosphorylation | فضا خارج سلولی، ماتریس خارج سلولی<br>Extracellular space, Extracellular matrix | اتصال یون روی، فعالیت اندوپپتیداز<br>Zinc ion binding, Endopeptidase activity |
| SLITRK4 | آکسونوژنسیس، تنظیم مثبت مجموعه سیناپس<br>Axonogenesis, Upregulation of synapse complex | غشاء<br>Membrane | اتصال به پروتئین<br>Protein binding |
| CHL1 | هدایت آکسون، تنظیم منفی فرایند آپوپتوز نورون<br>Axon guidance, Negative regulation of neuron apoptosis process | غشاء، دندریت، قسمت آپیکال سلول<br>Membrane, Dendrite, Apical part of the cell | اتصال پروتئاز، اتصال پروتئین<br>Protease binding, Protein binding |
| ZC3H12D | – | نوکلئوپلاسم<br>Nucleoplasm | اتصال یون فلزی<br>Metal ion binding |
| ADCY8 | فرآیند بیوسنتزی Camp، تنظیم پاسخ سلولی به استرس، هموستاز گلوکز، مسیر سیگنال دهی گیرندهG<br>Camp biosynthetic process, Regulation of cellular response to stress, Glucose homeostasis, G receptor signaling pathway | غشاء، دندریت و آکسون<br>Membrane, Dendrite and Axon | فعالیت آدنیلات سیکلاز، فعالیت فسفراکسیژن لیاز<br>Adenylate cyclase activity, Phosphoroxygen lyase activity |
| ENSGALG00000009041 | – | – | اتصال به پروتئین |
| RUNX1 | تنظیم رونویسی الگوی DNA، تنظیم تمایز سلولی<br>Regulation of DNA template transcription, Regulation of cell differentiation | هسته، نوکلئوپلاسم، سیتوپلاسم<br>Nucleus, Nucleoplasm, Cytoplasm | اتصال ATP، اتصال DNA، فعالیت فاکتور رونویسی متصل به DNA<br>ATP binding, DNA binding, DNA binding transcription factor activity |
| PHACTR3 | – | نوکلئوپلاسم<br>Nucleoplasm | اتصال پروتئین فسفاتاز۱، فعالیت تنظیم‌کننده پروتئین فسفاتاز نوع۱<br>Binding of protein phosphatase 1, Regulatory activity of protein phosphatase type 1 |
| ENSGALG00000008599 | فرایند بیوسنتز کربوهیدرات<br>Carbohydrate biosynthesis process | غشاء<br>Membrane | فعالیت سولفوترانسفراز<br>Sulfotransferase activity |
| FFAR4 | تمایز سلول‌های چربی، تنظیم انتقال گلوکز، ترشح هورمون، تنظیم منفی پاسخ التهابی<br>Differentiation of fat cells, Regulation of glucose transport, Hormone secretion, Negative regulation of inflammatory response | غشای پلاسمایی، وزیکول اندوسیتیک<br>Plasma membrane, Endocytic vesicle | اتصال اسید چرب، فعالیت گیرنده‌های همراه با پروتئین G<br>Fatty acid binding, Activity of G protein-coupled receptors |
| ENSGALG00000046947 | فاگوسیتوز، پاسخ ایمنی ذاتی<br>Phagocytosis, Innate immune response | غشای پلاسمایی، مجموعه ایمونوگلوبولین<br>Plasma membrane, Immunoglobulin complex | اتصال آنتی ژن، اتصال گیرنده ایمونوگلوبولین<br>Antigen binding, Immunoglobulin receptor binding |
| DGKB | انتقال سیگنال درون سلولی، فرایند متابولیک | غشاء<br>Membrane | اتصال یون کلسیم، اتصال ATP، فعالیت |

| | | | |
|---|---|---|---|
| | گلیسرولیپید<br>Intracellular signal transduction, Glycerolipid metabolic process | | دیاسیل گلیسرول کیناز<br>Calcium ion binding, ATP binding, Diacylglycerol kinase activity |
| ENSGALG00000051617 | پاسخ ایمنی ذاتی، فاگوسیتوز، مسیر سیگنالینگ گیرنده سلول B<br>Innate immune response, Phagocytosis, B cell receptor signaling pathway | مجموعه ایمونوگلوبولین، قسمت خارجی غشاء<br>Immunoglobulin complex, Outer part of the membrane | اتصال انتی ژن، اتصال گیرنده ایمونوگلوبولین<br>Antigen binding, Immunoglobulin receptor binding |
| KREMEN1 | مسیر سیگنال‌دهی Wnt<br>Wnt signaling pathway | غشاء<br>Membrane | – |
| CYP3A5 | فرآیند کاهش اکسیداسیون<br>Oxidation reduction process | – | غشاء<br>Membrane |
| CORIN | پروتئولیز، تنظیم دفع سدیم کلیوی<br>Proteolysis, Regulation of renal sodium excretion | غشاء، بدنه هسته، اسکلت سلولی اکتین<br>Membrane, Nuclear body, Actin cytoskeleton | فعالیت اندوپپتیداز، اتصال پروتئین<br>Endopeptidase activity, Protein binding |

در بررسی آنالیز هستی شناسی ژن، در مرحله شروع، از ژن‌های ENSGALG00000052817، ENSGALG00000049241 و ENSGALG00000051964 هیچ اطلاعاتی یافت نشد. در هر دو گروه، ژن‌های با بیشترین بیان در فعالیت‌های متابولیسم چربی و پاسخ ایمنی نقش دارند و در این مرحله آهکی شدن پوسته، ژن ENSGALG00000044418 با بیشترین تفاوت بیان، در فعالیت تنظیم سیستم خون‌رسانی و رگ‌زایی نقش دارد. به‌طور عمده ژن‌های شاخص شناسایی شده، جدول ۴، به‌ترتیب در فعالیت‌های اتصال پروتئین (ژن‌های ENSGAL00000044418، NMRAL1، SLITRK4، CHL1، ENSGAL00000009041، PHACTR3 و CORIN)، تنظیم رونویسی DNA (ژن‌های LHX1، TFAP2A، RUNX3، ETV1، NR4A3 و RUNX1)، پاسخ ایمنی (ژن‌های ENSGAL00000043996، TMEM117، AvBD9، GCNT3، ENSGAL00000046947 و ENSGAL00000051617)، متابولیسم چربی (ژن‌های PNPLA3، NR4A3، CD36، FFAR4 و DGKB) و اتصال یون کلسیم (ژن‌های ANXA10، CAPN8، CDH18 و DGKB) نقش دارند.

**جدول ۵**- آنالیز هستی‌شناسی ژن‌های شاخص متفاوت بیان شده در مرحله رشد دوره آهکی شدن پوسته

**Table 5**- Gene ontology analysis of indicator genes differentially expressed in the growth stage of the calcification period

| نام ژن<br>Gene name | فرآیند زیستی<br>Biological process | اجزای سلولی<br>Cellular components | عملکرد مولکولی<br>Molecular function |
|---|---|---|---|
| TMEM63C | حمل کاتیون<br>Cation transport | غشاء<br>Membrane | فعالبت کانال کاتیونی<br>Cation channel activity |

**جدول٦**- آنالیز هستی‌شناسی ژن‌های شاخص متفاوت بیان شده در مرحله خاتمه دوره آهکی شدن پوسته

**Table6**- Gene ontology analysis of indicator genes differentially expressed in the termination stage of the calcification period

| نام ژن<br>Gene name | فرآیند زیستی<br>Biological process | اجزای سلولی<br>Cellular components | عملکرد مولکولی<br>Molecular function |
|---|---|---|---|
| ENSGALG00000054226 | مسیر سیگنالینگ گیرنده همراه با پروتئین G، تشخیص محرک شیمیایی دخیل در درک حس بویایی<br>G protein-coupled receptor signaling pathway, Chemical stimulus recognition involved in olfactory perception | غشاء<br>Membrane | فعالیت گیرنده‌های بویایی<br>Activity of olfactory receptors |
| OGDH | چرخه تری کربوکسیلیک اسید، فرایند کاهش اکسیداسیون<br>Tricarboxylic acid cycle, Oxidation-reduction process | هسته، میتوکندری<br>Nucleus, Mitochondria | اکسی گلوترات دهیدروژناز، اتصال یون روی<br>Oxyglutarate dehydrogenase, Zinc ion binding |

در بر سی آنالیز ه ستی شنا سی ژن، در مرحله ر شد، از ژن‌های ENSGALG00000049618، ENSGALG00000048945 و در مرحله خاتمه، از ژن‌های ENSGALG00000049618، HSPB9 و MIR12247-1 اطلاعاتی یافت نشد.

در پژوهشی، با هدف بر سی و مقایسه مسیرهای ژنومی انتخاب مقایسه‌ای و ژن‌های کاندیدای کلیدی مرتبط با صفات اقتصادی در ژنوم نژادهای بومی مرغ روسی سفید نوع تخم‌گذار و مرغ کورنیش سفید نوع گوشتی در نهایت، ۴۵ منطقه ژنومی از جمله ژن‌های ANXA10، CD36 و PAPPA2 شناسایی شده‌اند (Abdelmanove et al., 2021).

در پژوه شی دیگر، با هدف بر سی دقت پیش‌بینی ژنومی صفت کیفیت پوسته تخم مرغ در لاین لگهورن سفید، ژن CAPN8 با نقش در فعل و انفعالات با یون کلسیم و نقش در آهکی سازی پو سته تخم مرغ مورد شناسایی گردیده است (Wolc et al., 2020).

در پژوه شی دیگر، پروفایل بیان ژن برای شنا سایی پروتئین‌های پوسته تخم مرغ که در دفاع فیزیکی آن نقش دارد مورد بررسی قرار گرفته است، که نتایج نشان داد، ژن AvBD9 در رحم مرغ بیان شده و پروتئین ضد میکروبی آن می‌تواند به محافظت در مقابل باکتری‌ها کمک کند (Jonchere et al., 2010).

## نتیجه‌گیری کلی

در آنالیز بیان افتراقی ژن، با در نظرگرفتن log2(fold_change) بیشتر از ۴+ و کمتر از ۴-، در مرحله شروع، ۵۱ ژن، در مرحله ر شد، دو ژن و در مرحله خاتمه چهار ژن مشخص شدند. بیشترین تفاوت بیان ژنی بین دو گروه در مرحله شروع، ژن‌های ENSGALG00000044418 و PDZK1IP1، در مرحله رشد، ژن‌های ENSGALG00000049618، ENSGALG00000048945 و TMEM63C و در مرحله خاتمه، ژن‌های ENSGALG00000049618 و ENSGALG00000054226 مشاهده شدند؛ نتایج آنالیز هستی‌شناسی ژن‌های شاخص نشان داد که به‌طور عمده به‌ترتیب در فعالیت‌های اتصال پروتئین، تنظیم رونویسی DNA، پاسخ ایمنی، متابولیسم چربی و اتصال یون کلسیم نقش دارند.

## References


1. Abdelmanova, A. S., Dotsev, A. V., Romanov, M. N., Stanishevskaya, O. I., Gladyr, E. A., Rodionov, A. N., Vetokh, A. N., Volkova, N. A., Fedorova, E. S., Gusev, I. V., Griffin, D. K., Brem, G., & Zinovieva, N. A. (2021). Unveiling comparative genomic trajectories of selection and key candidate genes in egg-type russian white and meat-type white cornish chickens. *Biology*, *10*(9), 876. DOI: 10.3390/biology10090876
2. Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. From: http://www.bioinformatics.babraham.ac.uk/projects/fastqc
3. Bahrami, A. (2020). Which aligner software is the best for our study?. *Genetics Research*, *7*, 048. DOI: 10.23937/2378-3648/1410048
4. Banabazi, M. H., Naserkheil, M., & Miraei-Ashtiani, R. (2012). Network regulatory gene expreesion of cerevisiae cell cycle with correlation weight method. The third national conference on agricultural biotechnology iran. Ferdowsi University of Mashhad, Iran. (In Persian).
5. Banabazi, M. H., Naserkheil, M., & Miraei-Ashtiani, R. (2012). Algorithm to identify genes expressed differently in the yeast *Saccharomyces cerevisiae* microarray data using software packages R. The third national conference on agricultural biotechnology Iran. Ferdowsi University of Mashhad, Iran. (In Persian).
6. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114-2120. DOI: 10.1093/bioinformatics/btu170
7. Ekblom, R., & Galindo, J. (2010). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, *107*, 1-15. DOI: 10.1038/hdy.2010.152
8. FAO. (2020). FAOSTAT. https://www.fao.org/faostat/en/#data/QCL/visualize/
9. Gan, Q., Chepelev, I., Wei, G., Tarayrah, L., Cui, K., Zhao, K., & Chen, X. (2010). Dynamic regulation of alternative splicing and chromatin structure in Drosophila gonads revealed by RNA-seq. *Cell Research*, *20* (7), 763-783. DOI: 10.1038/cr.2010.64
10. Github. (2017). Cufflinks. http://cole-trapnell-lab.github.io/cufflinks/
11. Haas, B. J., & Zody, M. C. (2010). Advancing RNA-seq analysis. *Nature Biotechnology*, *28*, 421. DOI: 10.1038/nbt0510-421
12. Jonchere, V., Rehault-Godbert, S., Hennequet-Antier, C., Cabau, C., Sibut, V., Cogburn, L. A., Nys, Y., & Gautron, J. (2010). Gene expression profiling to identify eggshell proteins involved in physical defense of the chicken egg. *BMC Genomic*, *11*, 57. DOI: 10.1186/1471-2164-11-57
13. Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*, 357-360. DOI: 10.1038/nmeth.3317
14. Marguerat, S., & Bahler, J. (2010). RNA-seq: From technology to biology. *Cellular and Molecular Life Sciences,*

*67*, 569-579. DOI: 10.1007/s00018-009-0180-6

15. Mortazavi, A., Williams, B., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods, 5*(7), 621-628. DOI: 10.1038/nmeth.1226
16. Sultan, M., Schulz, M. H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., & Schmidt, D. (2008). A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science, 321*(5891), 956-960. DOI: 10.1126/science.1160342
17. Wilhelm, B. T., & Landry, J. R. (2009). RNA-seq quantitative measurement of expression through massively parallel RNA-sequencing. *Methods*, *48*, 249-257. DOI: 10.1016/j.ymeth.2009.03.016
18. Wolc, A., Drobik-Czwarno, W., Jankowski, T., Arango, J., Settar, P., Fulton, J. E., Fernando, R. L., Garrick, D. J., & Dekkers, J. C. M. (2020). Accuracy of genomic prediction of shell quality in a white leghorn line. *Poultry Science, 99*, 2833-2840. DOI: 10.1016/j.psj.2020.01.019